Research Article

Prediction of Mortality Rate among ICU Patients based on Vital Signals using Data Analysis Methods

Mojtaba Johari Moghadam[1*], Nazanin Johari Moghadam[2] and Adel Joharimoghadam[3]

[1]School of Mechanical Engineering, Iran University of Science & Technology, Tehran, Iran; [2]MA Bio Medical Engineering, Islamic Azad University, Science and Research branch, Tehran, Iran; [3]Department of Cardiology, Aja University of Medical Sciences, Tehran, Iran

Corresponding author:
Mojtaba Johari Moghadam,
School of Mechanical Engineering,
Iran University of Science & Technology,
Tehran, Iran
Tel: 00989127110017;
E-mail: mojtabajoharimoghadam@yahoo.com

Abstract

Using a scoring system for illness intensity can be a signpost for the physician to objectively assess the future state of the patient or estimate the probability of his/her recovery. In fact, due to individual differences among ICU patients, it seems necessary to develop a special estimation system for every intensive care unit. This system helps physicians to prioritize the patients in receiving appropriate care and helps them to be more explicit in explaining and presenting the final state of the illness to the patients' affiliates. By analyzing the vital signals of ICU patients, their mortality rate can be automatically predicted. In our proposed system, using statistical analyses and Wavelet change, several features were extracted out of the patients' vital signals. Then using Fisher's standard analysis and multi-layered perceptron neural network classification, it was found that the proposed system is able to predict people's mortality rate with high accuracy. The results of the present study indicate that the proposed system has an accuracy level of 85.1% and 81.7% for treatment and test data, respectively. Any attempt to practically apply this method can help enhance the quality of care system in ICU sections of the hospitals considerably.

Keywords: Prediction; Mortality Rate; ICU; Vital Signals

Introduction

Intensive Care Unit or ICU is a specialized place in which the personnel and medical equipment are used to treat and manage extremely ill patients. One acceptable goal of setting up ICU is saving the lives of recoverable patients since all patients hospitalized in ICU do not return to their healthy or even normal lives, and some of them die as a result of the intensity of their illness. Since ICU is affected by different factors, offering appropriate care and treatment may have a considerable inhibiting effect on the illness. More than two million deaths occur in USA annually which can be reduced by 30% through appropriate services in ICU.

The ICU physicians also bear the responsibility of intensive care, identification, and separation of the ICU patients since, as already mentioned, all ICU patients do not necessarily benefit from this unit, and in the case of some ICU patients, it only leads to smoother and easier death of the patients. Using scoring system for illnesses can serve as a guide for the physician to objectively assess the future state of the patient or to estimate their probability of recovery. These pre-informative systems can also be helpful in estimating the physiological instability of the patients at their arrival to the hospital. Moreover, the intensity of illness can be scored, and mortality rate in each patient can be predicted while the patient is being clinically observed by ICU physician so that the probability of recovery may be assessed more accurately.[1]

Generalized linear model systems have already been widely used to predict the mortality probability of ICU patients. Three of the most commonly-used methods in this area include Acute Physiology and Chronic Health Evaluation (APACHE) called as APACHE scoring system,[2] Simplified Acute Physiology Score (SAPS),[3] and Mortality Probability Model (MPM).[4] In the last version of these systems, the worst physiological values are reported 24 hours after the patients are sent to ICU, and they are used for developing a linear predictive model.[5] Meanwhile, it seems that failure to use other sets of data and the loss of this type of information will yield inaccurate results about predicting the physiological state of the patients.

With the developments in computational technologies and high volume of clinical information in ICU, a growing number of researchers prefer to use data-based learning to predict mortality rate. Machine learning as a type of non-linear modelling methods have also been recently used in this area. For example, Artificial Neural Networks,[6] Vector Machine Support,[7] Decision Making Tree,[8] and Simple Bayesian Model [9] have been utilized to show patients' characteristics, and they have yielded much more accurate results. In all of these papers, it has been emphasized that with higher physiological information from the available patients, highly accurate results can be obtained.

In this article, the physiological information about ICU patients have been used to design a system that yields rather appropriate results for the physicians with the least possible amount of information so that the results obtained from mortality rate prediction model can be used in treating the patients.

How to Cite this Article: Moghadam MJ, et al. Prediction of Mortality Rate among ICU Patients based on Vital Signals using Data Analysis Methods. Ann Med Health Sci Res. 2019;9: 603- 607

Methodology

Extraction of the features

One of the most common methods of time series analysis and extraction of features from signals is the use of statistical parameters. The most important statistical features used in analyzing the signals includes mean, standard deviation, skewness, and kurtosis which are known as the first, second, third, and fourth moment of the data. In statistics, mean (mathematical mean) or average is a measure of central tendency, and it refers to the sum of the values in a data set divided by their frequency.

Standard deviation (which is usually shown by σ) is one of the dispersion indexes which shows the distance of the pieces of data from the mean. If standard deviation is considered to be a set of data near zero, then the data values are near the mean and have little dispersion while higher SD value indicates higher dispersion of scores. Standard deviation is the second root of variance. The advantage of using standard deviation instead of variance lies in the fact that standard deviation has the same dimension as that of the data.

Skewness shows the rate of asymmetry in probable distribution. If the data are symmetric with regard to the mean, the value of skewness will be zero. Skewness equals the third normalized moment, and it is, in fact, a criterion for symmetry/asymmetry of dispersion vector. For an asymmetric distribution with a tendency of the histogram towards higher values, the skewness value is positive, but for asymmetric distribution with a tendency towards lower values, the skewness value will be negative. Moreover, in statistics kurtosis describes the peakedness of a frequency distribution. The more peaked the shape of the histogram, the higher the kurtosis index rate for that piece of data would be and vice versa.

One of the other rather powerful tools for examining and analyzing the signals is Wavelet transform which has wide applications in the area of processing some features of the signals such as discontinuities, sudden local changes as well as the concentration of the energy in a small part of the signal. As for Fourier Transform, Wavelet transform is a powerful instrument in processing the signals which is necessary for transmitting signals from temporal space to another space. This is a three-dimensional space the dimensions of which include time, scale, and domain. The significance of this change lies in the fact that the 3-D space presenting this type of signal reveals some features of the signal which are not accessible with other signal processing tools.

Continuous Wavelet transform can be written as the internal multiplication of the signal and a base function in the following way:

$$(1) \quad CWT_x^{\varphi}(\tau,s) = \Psi_x^{\varphi}(\tau,s) = \frac{1}{\sqrt{|s|}} \int_{-\infty}^{+\infty} x(t)\, \Psi^*\left(\frac{t-\tau}{s}\right) dx$$

$$= \, < x(t),\ \varphi_{\tau,s}(t) >$$

In which: -

$$(2) \quad \varphi_{\tau,s}\ (\mathrm{t}) = \frac{1}{\sqrt{|s|}}\ \psi\left(\frac{t-\tau}{s}\right)$$

Based on the definition given in Equation (1) which is presented in the form of internal multiplication, it can be concluded that Wavelet change is, in fact, a measurement of the similarity between the signal and base functions (mother Wavelets). By the term similarity, we mean similarity assessment among the frequency-based content. In other words, Wavelet change coefficients indicate the extent of similarity of the signal to the Wavelet at the concerned scale.

Therefore, if the concerned signal is an outstanding index in the corresponding frequency with the scale under analysis, then the Wavelet being compared would be similar to the concerned signal. As a result, the coefficient value of changing continuous Wavelet obtained at this scale will be a relatively high value. Using the feature of base function transformation coefficients in the equation are calculated as follows:

$$(3) \quad f(t) = \sum_{k=1}^{N} a_k \phi_k(t) = \sum_{k=1}^{N} < \mathrm{f}(\mathrm{t}),\ \mathrm{f}_k(\mathrm{t}) > \phi_k(t)$$

Considering the role of computers in the computations today, it is necessary to both propose processing ideas and turn them into the data that can be calculated with computer. All of the changes mentioned so far are all continuous, and it is not possible to use them practically in computers. Therefore, it is necessary to use their discontinuous version. As for discontinuous Fourier Transform, in discontinuing Wavelet change the simplest procedure is sampling frequency time plate in its different positions.

Similarly, even sampling would be the simplest way in this regard. However, in the case of Wavelet change sampling rate can be reduced by changing the scale. Therefore, with high scale changes (lower frequencies) sampling rate may be reduced according to nicest rate.[10]

Classification

One of the major purposes of registering and processing vital signals is interpreting it and using its useful information in diagnosis and treatment. Interpretation occurs at recognition or classification stage. In fact, the classifying agent puts a label for each class in the features space. The present study has used artificial neural networks for data classification. Artificial neural networks are new systems and computation methods for machine learning.[11] The main idea behind such networks is demonstration of knowledge and its final application with the aim of predicting the output answers in complicated systems.

One simple family of neural networks is perceptron model in which there are n inputs and m outputs. For each input, there is a weighing coefficient of Wi, and for every output there is a threshold level labelled as q. In the topological structure of these networks, there is an input layer which receives the information; there are some hidden layers that receive the information from previous layers, and finally there is an output layer to which the results of calculations go and in which the replies are placed.

There are also single-layered and multi-layered networks. in which the single layer organization of all units are connected to one layer; these networks are most commonly used and have greater computational potential than multi-layered organization. In multi-layered networks, the layers are numbered by layers,

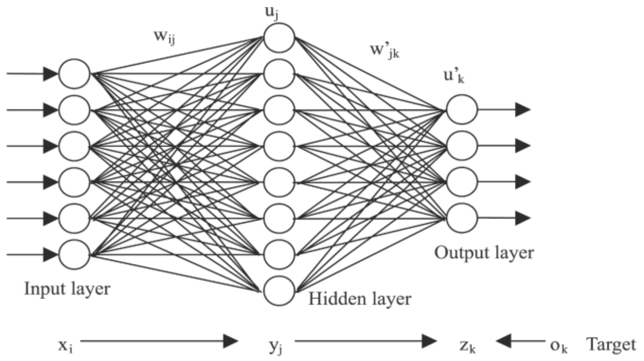

Figure 1: A multilayer perceptron neural network with a hidden layer.

and every two layers of the same network are connected to each other through weights or connections. These networks operate in the following way [Figure 1]:

- **Input layer:** Receiving the raw information fed into the network.
- **Hidden layers:** The performance of these layers is determined by the input and the weight of the connection between them and the hidden layers.
- **Output layer:** The performance of output unit depends on the activity of the hidden layer and the weight of connection between the hidden layer and the output.

Three parameters including sensitivity (Sensitivity-SE), Specificity rate (Specifity-SP), and accuracy (Accuracy-AC) are usually used for assessing classifying systems. Accuracy is a criterion to assess the rate of consistency between the obtained results and real results. The more experimental results are nearer to the real values, the more accurate they are. These parameters are computed in the following way:

$$\mathrm{SP} = \frac{TN}{(TN+FP)} \times 100 \quad AC = \frac{TP+TN}{TN+FP+FN+TP} \times 100 \quad \mathrm{SE} = \frac{TP}{(TP+FN)} \times 100$$

In the equations given above, TP refers to real positive cases, TN indicates real negative cases, FN refers to false negative cases, and FP indicates false positive cases.

- FP classifies the healthy under the label of patient.
- TP classifies the patient under the label of patient.
- FN classifies the patient under the label of healthy.
- TN classifies the healthy person under the label of healthy.

Results

Introducing the data

In the present study, the data presented in physio-net website concerning prediction of mortality rate was used.[12] In this base, the information about 4000 patients hospitalized in ICU section was collected out of which 3446 patients were released from hospital after passing through treatment phases; a total number of 544 patients did not survive (the proportion of mortality rate to the total population was 1 to 8).

The participants included 2246 male patients with a mean age of 62.9±17.2 and 1754 female patients with a mean age of 66.0±17.9. The patients were hospitalized for different reasons including: suffering from cardio-vascular illnesses (577 people), recovery from heart surgery (874 people), and other urgent illnesses (1481 people), and recovery from surgeries other than heart surgery (1068 people). Totally 37 cases of vital signals were reported from different people at the time of hospitalization. It should be noted that totally 37 cases have been registered, and for one or more people, some of the values have not been registered.

Pre-processing the data

After initial data analysis, it was found that definitely there is a primary analysis of the data. Therefore, the following features were taken into account, and in case a piece of data didn't have these features, it was eliminated.

Out of 37 registered vital variables, initially the parameters which were collected for more than 95% of the people were maintained, and the others were eliminated. In this way, 18 parameters out of a total number of 37 were removed. Then, Glasgow Coma Scale (GCS) was used to study 19 variables were examined including: age, gender, height, the reason for hospitalization, weight, the rate of heart beat, Diastole blood pressure, the time and amount of urination at the time of hospitalization, the nitrogen in blood urine (BUN), the creatinine disposed by the kidneys, the rate of glucose in the blood, the rate of Bicarbonate and Hematocrit in the blood, hypokalaemia, the percentage of Magnesium and Sodium, the percentage of platelets, the number of white globules in a unit of blood, body temperature, and consciousness.

Then for any of the people for whom one or more of the above-mentioned 19 variables have not been registered were removed again. In this way, 269 people including 26 of the patients who had died, and 243 people who checked out of the hospital were eliminated leaving a total number of 3731 people (3203 patients who had checked out and 528 people who had died).

One of the most important parameters in this kind of research is the rate of heart beat (HR). Therefore, all data which registered less than 5 HRs were eliminated. The reason lies in the fact that after dramatic decrease in the number of registered HRs, it was not feasible to conduct some analyses including Wavelet analysis. Therefore, 14 other pieces of data including 12 released patients and 2 dead ones were eliminated, and finally 3713 pieces of data (3191 patients who were released and 526 dead patients) were left.

Extraction of features

The following features were used in the present study:

- The values obtained for age, gender, height, and reason for hospitalization (4 features)
- For the other 15 variables, the values of mean, minimum, maximum, standard deviation, and ascending or descending slope were considered as features (15*5=75 features),
- According to the reference values, the values for skewness and kurtosis for HR rate, the rate of urination, and body temperature were extracted as features (6=3*2 features).
- The standard deviation for the signals obtained from changing the Wavelet of HR rate was also considered as a feature. To do so, the main signal was divided into two types of detail signals (high frequency) and approximate

signals (low frequency) using one phase of Wavelet change, and the standard deviation for each one was considered as a feature (2 features).

The proportion of HB frequency with a difference of more than 50 milliseconds to the total number of beats at the time of hospitalization was extracted using kubios software and was considered as a feature.

Therefore, 88 features have been extracted using the above mentioned methods which will be used later. It is clear that all of these features are not useful and valid for identifying and predicting mortality. Therefore, it is necessary to eliminate some of the best criteria using a single criterion. Fisher's criterion was used to rank the features in the present study.

Fisher's criterion is the distinction which every feature can establish between two groups. The equation for Fisher's criterion for each feature is as follows; in this equation u1 and u2 are the mean of the data for each feature in the first and second group; s1 and s2 are the standard deviation of the data for each feature of two classes of the data. The value of Fisher's criterion for all computed features has been calculated and drawn in Figure 2.

$$F(W) = \frac{(u_1 - u_2)^2}{\left(S_1^{\,2} + S_2^{\,2}\right)}$$

For exact determination of the number of usable features, they were arranged according to the obtained value of Fisher's criterion. In this way, the feature with the highest value for Fisher's criterion value (Feature 28) was ranked as first, and the feature with the lowers Fischer criterion was ranked as the last. The number of the features started with one, and classification was conducted in a way that in each stage one feature was added.

In all of these steps, the available data were classified into two groups of neural network trainers (70%) and test data (30%). The important point to note is the fact that the number of the people who checked out of the hospital was 7 times as many as the dead patients. In this way, the network allots greater importance to the data from the healthy people. Therefore, in order to train the network, the pieces of data in both groups should be equal in number.

In this line, as many patients as the patients who had died (after leaving hospital) were randomly selected. The effect of increasing the number of features on accuracy is shown in

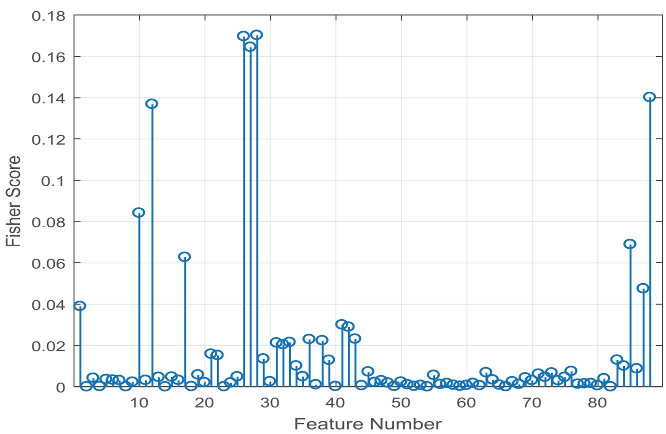

Figure 2: Computation of Fischer's criterion for all extracted features.

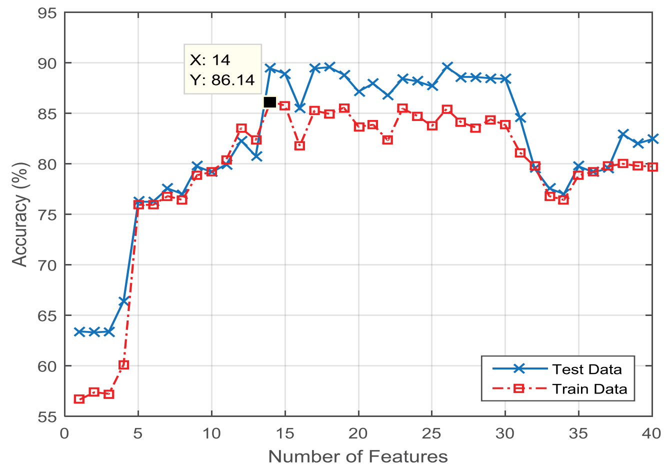

Figure 3: The effect of increasing the number of features based on the order of Fisher's criterion on classification accuracy.

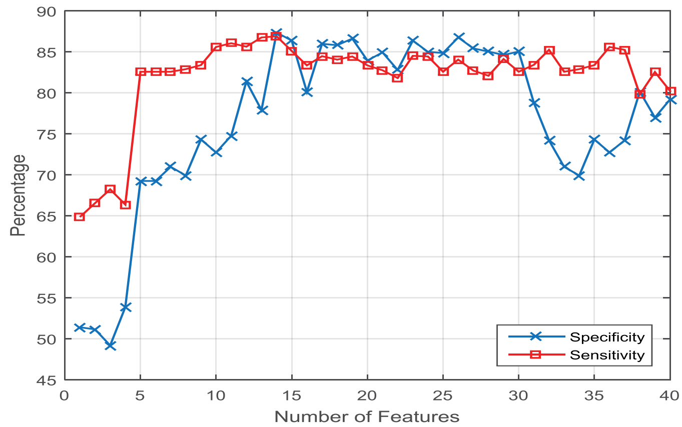

Figure 4: The effect of increasing the number of features based on Fischer criterion on sensitivity and specialization.

Figure 3. Moreover, Figure 4 shows the rate of sensitivity and specialization for the test data with an increase in the number of features.

Discussion

It was explained in the previous section that after preprocessing the data adopted from Physio-net website, the features were extracted from the signals. Pre-processing is a necessary issue since there have been many cases in which, due to human error or the short stay of the patient in the hospital, some of the most commonly used vital signals of the patients have not been recorded. Therefore, some values are missing and they should be eliminated. After pre-processing, the dimension of the data was reduced to 3717*19 in which 3717 shows the number of the participants in the experiment (including 3191 checked-out patient and 526 dead patients). There are also 19 vital signals which have been mentioned in section 2.3.

After conducting pre-processing, using the available signals, 88 features were extracted for each person most of which were obtained through statistical analysis of the signals; some features were also extracted based on Wavelet change.

Since HR signal has gained greater significance in research than the other types, the present study has also focused on this type of signals. Although these signals clearly have relatively good results in the accuracy of predicting mortality rate, they suffer from some disadvantages the most important of which is unfeasibility of hardware application due to many computations involved and over consuming time due to high amounts of data.

On the other hand, there is not guarantee to assure that all of the mentioned 88 features have had an effective role in predicting mortality rate. Therefore, Fischer's criterion was used to rank the number of features, and the obtained results indicated that only 14 features are actively involved in this issue.

A shown in Figure 3, the greatest rate of accuracy for the test data (which are of greater importance than the others) among 14 features amounts to 86.1%. Figure 4 also shows clearly that the rate of sensitivity and specialization for the mentioned 14 features in the test data is maximal. Further examination revealed that for the parameters of age, HR, the amount of nitrogen in blood urine, the rate of bicarbonates dissolved in blood and the rate of consciousness only Glascow coma criterion has been used.

Therefore, the present study has used only 16 vital indexes and 14 features extracted from them. The list of favorable features obtained through adopted criterion is provided below:

- Mean Age of the People
- Mean HR
- Maximum HR
- Standard deviation of HR
- The percentage of pNN50
- Standard Deviation of Approximate value of HR Wavelet signal
- Mean rate of Nitrogen in blood urine
- Minimum Rate of Nitrogen in Blood Urine
- Maximum rate of nitrogen in blood urine
- Mean blood sugar
- Mean Rate of Bicarbonates in the blood
- Mean rate of bicarbonates in the blood
- Maximum consciousness based on Glascow coma criterion
- The slope of consciousness parameter based on Glascow coma criterion

The model proposed in the present study included a classification of multilayered perceptron neural network with a hidden layer that also included 20 neurons and finally led to the use of non-linear features of these networks and higher accuracy. After training the neural network, the results of the present study indicated that the designed plan was efficient for the test and training data with an accuracy level of 86.1% and 89.7%, respectively. It is natural that lower accuracy for the test data is natural since the neural network has the final answer (release or death) for the training data, and it tries to achieve a favorable answer by regulating its coefficients, but for the test data, the network does not have the desirable result at all.

Conclusion

The results of the present study have good consistency with the results of other studies. For example, in a study conducted recently by Ding et al., 85.6% was introduced as the highest accuracy. However, the database of the mentioned study is different from the data base used in the present study. Meanwhile, the results of both studies are almost similar to each other. Using the instrument introduced in this study which depended on only six parameters of age, HR, BUN, blood sugar, Bicarbonates, and CGS, it is possible to intervene in the treatment process of the patients and provide necessary timely cautions for the medical team. The results of the present study showed that by using a set of statistical features obtained from Wavelet change, mortality among the ICU patients can be predicted with high accuracy.

Conflict of Interest

The authors disclose that they have no conflicts of interest.

References

1. Llompart-Pou JA, Chico-Fernández M, Sánchez-Casado M, Salaberria-Udabe R, Carbayo-Górriz C, Guerrero-López F, et al. Scoring severity in trauma: comparison of prehospital scoring systems in trauma ICU patients. Eur J Trauma Emerg Surg. 2017;43:351-357.
2. Knaus WA, Draper EA, Wagner DP, Zimmerman JE. APACHE II: A severity of disease classification system. Crit Care Med. 1986;14:754-755.
3. Le GJ, Lemeshow S, Saulnier F. A new simplified acute physiology score (SAPS II) based on a European/North American multicenter study. JAMA 1993;270:2957-2963.
4. Lemeshow S, Teres D, Klar J, Avrunin JS, Gehlbach SH, Rapoport J. Mortality probability models (MPM II) based on an international cohort of intensive care unit patients. JAMA: 1993;270:2478-2486.
5. Ding Y, Wang Y, Zhou D. Mortality prediction for ICU patients combining just-in-time learning and extreme learning machine. Neurocomputing 2017.
6. Hussain AJ, Fergus P, Al-Askar H, Al-Jumeily D, Jager F. Dynamic neural net- work architecture inspired by the immune algorithm to predict preterm deliveries in pregnant women. Neurocomputing. 2015;151:963-974.
7. Chen P, Yuan L, He Y, Luo S. An improved SVM classifier based on double chains quantum genetic algorithm and its application in analogue circuit diagnosis. Neurocomputing 2016;211:202-211 .
8. Azar AT, El-Metwally SM. Decision tree classifiers for automated medical diagnosis. Neural Comput. Appl. 2013;23:2387-2403 .
9. Ryynänen OP, Soini EJ, Lindqvist A, Kilpeläinen M, Laitinen T. Bayesian pre- dictors of very poor health related quality of life and mortality in patients with COPD. BMC Med. Inform. Decis. Mak. 2013;13:1-10.
10. Cui Z, Wang Y, Gao X, Li J, Zheng Y. Multispectral image classification based on improved weighted MRF Bayesian. Neurocomputing 2016;212:75-87.
11. Kim KJ, Cho SB. Prediction of colon cancer using an evolutionary neural network. Neurocomputing 2004;61:361-379.
12. https://physionet.org/challenge/2012/4.